单身与否对生活满意度的影响——基于知乎数据的研究

张志勇[1,2] 马炜玮[1,2] 张诗雨[1,2] 邱 义[1,2] 徐成龙[1,2] 张洪雨[1,2] 朱廷劭[1,2] 赵 楠[1,2]

([1]中国科学院心理研究所，北京，100101)

([2]中国科学院大学心理学系，北京，100049)

摘 要 为探究单身人群和婚恋人群在生活满意度和情绪词词频比例上的差异，本研究依托网络大数据，使用 python 爬虫程序对知乎平台上单身主题和婚恋主题下的用户发言进行爬取，并通过“文心”系统和生活满意度预测模型获得对应人群的情绪词频比例和生活满意度得分。结果发现婚恋组的生活满意度显著高于单身组（t=4.415，p＜0.001）；婚恋组的正性情绪词（t=-9.061，p＜0.001）和焦虑词（t=1.844，p＜0.001）的比例显著低于单身组，愤怒词（t=5.101，p＜0.001）的词频比例显著高于单身组。结果表明，婚恋人群在获得更高生活满意度和更低焦虑水平的同时，也需要处理与伴侣相关的情绪问题。

关键词 单身人群；婚恋人群；词频分析；生活满意度；机器学习

1 引言

自 2013 年以来，我国结婚登记人数逐年减少，越来越多的适龄男女或主动或被动地选择保持单身状态。截至 2021 年我国单身群体约为 3 亿，根据预测未来十年单身人数仍将不断增加。单身人群已经成为社会中的重要组成部分，单身人群的生活状态、生活满意度亦是影响社会稳定、和谐的重要因素。2021 年 2 月 15 日央视报道最新调查结果表明，单身人群幸福感和生活满意度逆袭已婚人群，这无疑是对单身人群生活满意度较低——这一主流调查研究结果的一大挑战。目前已有研究探讨了单身与婚恋人群生活满意度差异，但尚未得出一致结论，随着近年来人们婚恋价值观的不断改变，越来越多的研究发现单身人群并没有以往想象中的那么不快乐（Chopik 等，2021），单身人群内的不同类型也得到了进一步的细分，包括性别、文化、人格特质等。Verbrugge（1979）认为与未婚人群相比，已婚人群拥有较高的幸福感和较低的焦虑水平；一项由 553 名 20~30 岁的青年男女参与的研究发现，是否单身对生活满意度无直接影响，但单身人群具有的更高的孤独感与更低的社会支持可能会间接损害其生活满意度（Adamczyk, & Segrin，2015）。而另外的一些研究得

出了不同的结论，Girme 等（2016）的研究结果表明高回避目标的单身人群生活满意度与婚恋人群相同；女性的生活满意度不受单身与否的影响，香港已婚女性的生活满意度低于未婚女性（Markey 等，2005；Yang, & Yen，2014）。

生活满意度是个体在主观上对其生活质量的综合评价，这种评价可以更为完善的反映个体生活状况在认知层面的评价，表现方式大多通过问卷的方式体现。生活满意度会受到多种主客观因素的影响，主观因素主要包括人际交往、人格特质、价值观等认知和情感因素(Veenhoven，1984)，客观因素的研究则主要集中在收入、文化、生活环境、社会政策和福利等方面（朱迪，2016）。

单身与否对人们生活满意度的影响及其影响机制存在差异。一方面，恋爱关系具有“选择效应”和“保护效应”。“选择效应”是指处于恋爱关系的双方往往是由于某些特征相互吸引而结合在一起，与自己情投意合的人相处会提高彼此的幸福感；“保护效应”是指因为恋爱不仅有利于心理和身体健康，也使人们更容易走出生活困境，得到情感和经济方面的支持和帮助，有助于增加生活满意度（高夏丽，2021）。而另一方面，相较于婚恋人士会在很大程度上因受恋爱关系影响减少其他的社交活动，单身人士有更多的时间和家人、朋友、同事相处，从多元化的关系中获得更高的幸福感。主动选择单身的人群往往经济优越、学历较高、有一定的社会地位，且具有更高的自我认同感，具有这些特性的个体有足够的收入和能力来满足自己的欲望与需求，追求自我价值的实现，从而更有可能拥有较高的生活满意度（宋健, & 王记文，2016）。

基于上述理论背景以及不同研究的矛盾结果，本研究试图从一种新的方法学视角，利用网络数据平台，旨在通过用户网络发言表述，探讨单身与婚恋人群的生活满意度差异，并尝试对其背后的原因进一步进行分析。研究青年群体单身与婚恋这两种状态与生活满意度之间的关系，不仅对适婚人群树立正确的恋爱观、婚姻观具有积极导向作用，也为国家制定婚恋政策提供理论及数据参考，从而引导人们构建稳定的婚恋关系。本研究在为社会学家、人口学家、心理学家的研究提供参考资料的同时，亦开拓了使用社交媒体数据这一新视角。

2 方法

2.1 被试选取

本研究数据来源于知乎平台用户在相关话题下的回答。知乎是目前中文互联网上最大的深度内容聚集地，用户可以在其感兴趣的问题下进行回答、讨论。选取知乎回答比较单身人群和婚恋人群生活满意度具有以下优势。

与传统问卷研究相比采用社交媒体研究具有突出优势：一是生态效度高，知乎允许用户匿名回答，排除问卷研究中社会赞许性的影响，数据真实性高；二是相比问卷研究获取的样本量大且成本较低；知乎平台获取的文本数据有效性高：知乎允许用户以回答的形式围绕某一主题进行描述，通过合理选取问题，提取目标群体相对容易，且这种形式使用户在相对理性的状态下进行全面思考，其文字描述具有客观性和全面性，绝大部分回答能切合生活状态描述的要求，与所研究的问题高度相关。另外，借助知乎问题描述和用户账户信息，可一定程度平衡被试的性别、职业等变量。

本研究在单身或婚恋人群生活状态自我描述有关的问题下对相关回答进行爬取，所获取的数据均公开可用，不涉及用户身份信息。

2.2 网络数据爬取

针对所选取的知乎提问，利用 python 程序模拟浏览器登录状态，请求知乎 API 链接，逐一抓取该提问下的回答数据，进行筛选并保存。

2.3 数据清洗

数据清洗包括程序清洗和人工清洗两个环节。

2.3.1 程序清洗

使用爬虫程序爬取单身主题和婚恋主题下的用户发言各约 950 条。爬取过程中对获取的每一条回答进行初筛，仅保留其中大于等于 300 字的回答文本。除此之外，利用程序去除带有外部链接的文本和带有“公众号”字样的文本，以剔除大部分广告或营销号、机构号发布的不能反映个人生活状态的回答。最后通过比较前后两个回答文本之间前 50 字是否相同，

剔除大量由于知乎用户更新回答而重复发布的文本信息，仅保留该用户回答的首发文本。

2.3.2 人工清洗

数据经程序清洗后要进一步经过人工筛选保留合格数据。人工筛选的标准包括：（1）回答切合主题：剔除完全文不对题的回答；此外，部分单身人士会在婚恋话题下描述自己的单身生活并明确表明单身身份，或反之明确表明自己身份的婚恋人士在单身话题下描述自己的婚恋生活，这种情况下则需要将这类数据重新正确归类；（2）剔除程序未能识别的营销号和机构号的回答；（3）剔除对某一现象进行评论而并非自我描述的回答；（4）剔除以小说等形式呈现的虚构回答；（5）剔除旁观者视角的文本；（6）剔除无意义文本。按上述标准筛选后，最终获得有效数据 1739 条，其中单身人群自我描述 861 条；婚恋人群自我描述 878 条。

2.4 内容特征提取

“文心”中文心理分析系统（TextMind）由中国科学院心理研究所计算网络心理实验室研发，包括中文分词工具和心理分析词典，以分析文本中使用的不同类别语言的程度、偏好等特点。本研究使用“文心”中文心理分析系统对清洗后的数据进行内容特征提取，利用“文心”系统中的简体中文 LIWC 词典计算单身人群和婚恋人群的情绪词词频比例。

2.5 生活满意度预测模型

生活满意度反映了个体对自身状态的主观满意程度。本研究使用志愿者的生活描述文本和生活满意度问卷评分作为标签数据，使用机器学习算法建立生活满意度预测模型，将清洗后的数据导入生活满意度预测模型以获得相应的生活满意度评分。

3 结果

3.1 词频分析

使用 SPSS26.0 对两组样本情绪词词频比例上的差异进行独立样本 T 检验。结果如表 1 所示，婚恋组的正性情绪词（t=-9.061，p＜0.001）和焦虑词（t=1.844，p＜0.001）的比例显著低于单身组，愤怒词（t=5.101，p＜0.001）的词频比例显著高于单身组。

表 1 词频分析描述性数据与独立样本 t 检验

情绪词	分组	M±SD	t	df	p
正性情绪词	婚恋组	0.028±0.012	-9.061	1737.000	0.000***
	单身组	0.034±0.016			
负性情绪词	婚恋组	0.017±0.008	1.844	1737.000	0.065
	单身组	0.017±0.010			
焦虑	婚恋组	0.002±0.002	-4.638	1737.000	0.000***
	单身组	0.003±0.004			
愤怒	婚恋组	0.004±0.004	5.101	1737.000	0.000***
	单身组	0.003±0.004			
悲伤	婚恋组	0.003±0.003	-0.683	1737.000	0.495
	单身组	0.003±0.004			

注：$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

3.2 生活满意度指标

使用独立样本 t 检验对生活满意度预测模型处理后的婚恋组和单身组的生活满意度数据进行比较，结果如表 2 所示，婚恋组的生活满意度显著高于单身组（t=4.415，p＜0.001）。

表 2 生活满意度描述性数据与独立样本 t 检验

生活满意度	M±SD	t	df	p
婚恋组	0.613±0.064	4.415	1737.000	0.000***
单身组	0.598±0.072			

注：$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$

4 讨论与展望

本研究为探究单身与否对个体生活满意度的影响，使用 python 爬虫程序对知乎上单身主题和婚恋主题下的用户发言进行爬取，并通过选取镜像问题的方式在一定程度上降低样本的选择偏差。对爬取的数据进行人工数据清洗，在原有的 1936 条数据下筛选出有效数据 1739 条，其中单身人群文本 861 条，非单身人群文本 878 条。利用文心系统和生活满意度预测模

型获得词频特征及生活满意度指标并进行组间比较，结果发现婚恋组的生活满意度显著高于单身组，词频分析结果显示婚恋组的正性情绪词和焦虑词的比例显著低于单身组，愤怒词词频比例显著高于单身组。

婚恋组的生活满意度显著高于单身组这一研究结果，符合本研究预期，与以往部分研究的结果一致。一方面，单身群体一定程度上面临着更大的经济压力及社会压力，同时生活中的孤独感也会产生巨大的精神压力，因此我们推测正是在这样多重的内外压力情况下，单身群体心理受到一定损耗，从而导致相对较低生活满意度。另一方面，婚姻对人们有很多潜在的好处，例如孩子、稳定的性生活、家庭和经济支持、年老时的伴侣等，婚姻带来的社会支持有利于身心健康，可提高主观幸福感（袁正，& 李玲，2017），同时婚姻具有“选择效应”和“保护效应”，也使人们更容易走出生活困境，得到情感和经济方面的帮助，有助于增加幸福感（高夏丽，2021），从而使婚恋群体具有更高的生活满意度。词频分析结果显示，婚恋组的正性情绪词和焦虑词的比例显著低于单身组，愤怒词词频比例显著高于单身组，这一结果看似与生活满意度提高存在矛盾的地方，但 Diener（2006）将“主观幸福感”分解为三个主要组成部分: (1) 生活满意度，指对生活总体的评价以及对生活具体方面，比如婚姻、工作、收入、住房和休闲的评价; (2) 情绪，包括正面的情绪如愉悦、满足和负面的情绪如沮丧、愤怒; (3) 价值判断，如对自己的生活是否抱有成就感，觉得人生是否有意义。由此可见，选择进入恋爱或者婚姻亲密关系，有助于提升生活满意度，但对于主观幸福感的情绪部分则有益亦有害。

本研究使用知乎平台爬取数据并进行分析，有生态效度高、数据量大、成本低等优势。但本研究有一定局限性。首先，单身群体相较非单身群体，生活相对自由，有大量的时间和精力，因此可能更多的使用网络平台，发表自我描述，这一差异是否会对我们的结果产生影响尚不明确有待深入研究。其次，单身群体选择单身可能出于不同原因，如一些人因为生活压力、经济压力过大被迫单身，而另一些人则是更加注重事业及自我实现，不同动机的单身群体生活满意度应该会有所差异，这也需要再进行深入研究。再次，本研究词频分析结果表明，婚恋组愤怒词词频比例显著高于单身组，正性情绪词词频比例显著低于单身人群，这反应出婚恋也会带来一定的负面影响，但这一差异原因有待进一步研究分析其原因。最后，本研究仅区分了单身与非单身两个群体，群体内更加细致的划分，如性别、年龄、受教育程度等，还没有充分挖掘知乎平台所能够提供的公开用户的这些信息，下一步可以充分利用信息对数据的相关变量进行进一步控制，以确定对生活满意度的影响。

5 结论

本研究选择知乎作为数据来源，使用 python 爬虫程序爬取相关话题下的用户发言，经过程序筛选和人工筛选后获得对应人群的文本数据并通过“文心”系统和生活满意度预测模型获得对应人群的情绪词频比例和生活满意度得分。经分析后发现，婚恋群体相较于单身群体具有更高的生活满意度和更低的焦虑水平，但是本研究的结果也显示婚恋群体的正性情绪词的词频比例显著低于单身群体，愤怒词的词频比例显著高于单身群体。

综上所述，婚恋人群在获得更高生活满意度和更低焦虑水平的同时，也需要处理与伴侣相关的情绪问题，本文为年轻群体的婚恋抉择提供了一定程度的建议。

参 考 文 献

朱迪.(2016).市场竞争、集体消费与环境质量——城镇居民生活满意度及其影响因素分析. 社会学研究(03),193-217+246. doi:10.19934/j.cnki.shxyj.2016.03.009.

高夏丽. (2021). 个体化视角下不婚女性的情感态度及情感调适研究. (2019-4), 100-112.

宋健, & 王记文. (2016). 中国青年的婚姻状态与主观幸福感. 中国青年研究(9).

高夏丽. (2021). 个体化视角下不婚女性的情感态度及情感调适研究. (2019-4), 100-112.

袁正 & 李玲.(2017).婚姻与幸福感:基于 WVS 的中国微观数据. 中国经济问题(01),24-35. doi:10.19365/j.issn1000-4181.2017.01.03.

Verbrugge, L. M. (1979). Marital status and health. Journal of Marriage and the Family, 267-285.

Adamczyk, K., & Segrin, C. (2015). Direct and indirect effects of young adults’ relationship status on life satisfaction through loneliness and perceived social support. Psychologica Belgica, 55(4), 196.

Girme, Y. U., Overall, N. C., Faingataa, S., & Sibley, C. G. (2016). Happily single: The link between relationship status and well-being depends on avoidance and approach social goals. Social Psychological and Personality Science, 7(2), 122-130.

Oh, J., Chopik, W. J., & Lucas, R. E. (2021). Happiness Singled Out: Bidirectional Associations Between Singlehood and Life Satisfaction. Personality and Social Psychology Bulletin.

Veenhoven, R. . (1984). Conditions of happiness. D. Reidel Pub.

Markey, C. N., Markey, P. N., Schneider, C., & Brownlee, S. (2005). Marital status and health beliefs: Different relations for men and women. Sex Roles, 53(5), 443–451.

Yang, W. S., & Yen, P. C. (2014). Single and happier? A comparative study of marital status, socioeconomic security and wellbeing of women in Hong Kong and Taiwan. In The family and social change in Chinese societies (pp. 81-97). Springer, Dordrecht.

Diener，Ed 2006，“Guidelines for National Indicators of Subjective Well-Being and Ill-Being．” Journal of Happiness Studies 7．

The impact of being single or not on life satisfaction - a study based on Zhihu data

ZHANG Zhiyong[1,2]MA Weiwei[1,2]ZHANG Shiyu[1,2]QIU Yi[1,2]XU Chenglong[1,2]ZHANG Hongyu[1,2]ZHU Tingshao[1,2]ZHAO Nan[1,2]

([1]Institute of Psychology, Chinese Academy of Sciences, Beijing 100101)

([2]University of Chinese Academy of Sciences, Beijing 10049)

Abstract：In order to explore the differences in life satisfaction and emotional word frequency between single people and married people, this study uses Python to collect data from Zhihu, China’s largest Q&A platform, and obtain the corresponding group's sentiment frequency ratio through the "Wenxin" system and life satisfaction score by a life satisfaction prediction model. Results showed that the life satisfaction score in the married group was significantly higher than that in the single group ($t=4.415$, $p<0.001$); The proportion of positive emotion words ($t=-9.061$, $p<0.001$) and anxiety words ($t=1.844$, $p<0.001$) in the married group was significantly lower than that in the single group, but the proportion of anger words ($t=5.101$, $p<0.001$) was significantly higher than that in the single group. The results show that while married people obtain higher life satisfaction and lower anxiety level, they also need to deal with partner related emotional problems.

作者贡献声明：

张志勇：提出研究思路，设计研究方法，数据分析，论文研究方法和结果、结论部分撰写，最终论文修订。

马炜玮：提出研究思路，设计研究方法，知乎数据爬取、清洗与整理，论文研究方法部分撰

写。

张诗雨：提出研究思路，设计研究方法，数据分析，论文引言部分撰写。

邱义：提出研究思路，设计研究方法，论文讨论与展望部分撰写。

徐成龙：提出研究思路，设计研究方法，论文中英文摘要撰写。

张洪雨：提出研究思路，设计研究方法，论文引言部分撰写。

赵楠、朱廷劭：技术分析指导、论文指导。